HKUST – Department of Computer Science and Engineering
COMP 2711: Discrete Math Tools for CS – Spring 2012

#

Final Examination

Date: 25 May 2012 Time: 16:30–19:30 Venue: HALL

Name: ______________ Student ID: ______________
Email: ______________ Lecture and Tutorial: ________

Instructions

- This is a closed book exam. It consists of 9 questions and 19 pages.
- Please write your name, student ID, email, lecture section and tutorial on this page.
- For each subsequent page, please write your student ID at the top of the page in the space provided.
- Please sign the honor code statement on page 2.
- Answer all the questions within the space provided on the examination paper. You may use the back of the pages for your rough work. The last three pages are scrap paper and may also be used for rough work. Each question is on a separate page. This is for clarity and is not meant to imply that each question requires a full page answer. Many can be answered using only a few lines.
- **Unless otherwise specified you *must* always explain how you derived your answer. A number without an explanation will be considered an incorrect answer.**

Questions	1	2	3	4	5	6	7	8	9	Total
Points	10	11	10	10	12	10	15	10	12	100
Score										

Student ID: ______________

As part of HKUST's introduction of an honor code, the HKUST Senate has recommended that all students be asked to sign a brief declaration printed on examination answer books that their answers are their own work, and that they are aware of the regulations relating to academic integrity. Following this, please read and sign the declaration below.

```
I declare that the answers submitted for
this examination are my own work.

I understand that sanctions will be
imposed, if I am found to have violated the
University regulations governing academic
integrity.

Student's Name: ______________________________

Student's Signature: _________________________
```

Student ID: ______________

Problem 1: [10 points]

1. Consider the recurrence below defined for $n \geq 0$.

$$T(n) = \begin{cases} 1 & \text{if } n = 0 \\ 4T(n-1) + 6 & \text{if } n > 0 \end{cases}$$

   Give a closed-form, exact solution to the recurrence. You only have to give the solution. You do **not** need to show how you derived it.

2. Prove the correctness of your solution by **induction**.

Solution:

1. By iterating the recurrence, we get:

$$T(n) = 4^n + 6\frac{4^n - 1}{4 - 1} = 3 \cdot 4^n - 2.$$

2. We need to prove:

$$T(n) = 3 \cdot 4^n - 2. \qquad (1)$$

   Base case: When $n = 1$, $T(0) = 1 = 3 \cdot 4^0 - 2$. Equation (1) is true.

   Inductive step: Now consider $n > 1$. Assume Equation (1) is true for the case of $n - 1$, i.e.,

$$T(n-1) = 3 \cdot 4^{n-1} - 2.$$

   For the case of n, we have

$$\begin{aligned} T(n) &= 4T(n-1) + 6 \\ &= 4(3 \cdot 4^{n-1} - 2) + 6 \quad \text{(induction hypothesis)} \\ &= 3 \cdot 4^n - 2. \end{aligned}$$

   By the weak principle of mathematical induction, we conclude that Equation (1) is true for all $n \geq 0$.

Grading: 5, 5 .

Student ID: ______________

Problem 2: [11 points] Let a be a non-negative real number and n be a non-negative integer that is a power of 3. Consider a function $T(n)$ given by

$$T(n) = \begin{cases} 1 & \text{if } n = 1 \\ aT(\frac{n}{3}) + n^2 & \text{if } n > 1 \end{cases}$$

1. Find a closed-form expression for $T(n)$ by iterating the recurrence. Show the steps.
2. For which values of a is each of the following equations true?
   (i) $T(n) = O(n^2)$?
   (ii) $T(n) = O(n^2 \log n)$?
   (iii) $T(n) = O(n^3)$?

   For this part, it is **not** necessary for you to justify your answers.

Solution:

1. Let $n = 3^h$ for some integer h. If $h > 1$, we have

$$\begin{aligned}
T(n) &= T(3^h) \\
&= aT(3^{h-1}) + 3^{2h} \\
&= a(aT(3^{h-2}) + 3^{2(h-1)} + 3^{2h} \\
&= a^2T(3^{h-2}) + a3^{2(h-1)} + 3^{2h} \\
&\vdots \\
&= a^hT(3^{h-h}) + a^{h-1}3^2 + a^{h-2}3^4 + \ldots + a3^{2(h-1)} + 3^{2h} \\
&= a^h(1 + \frac{9}{a} + (\frac{9}{a})^2 + \ldots + (\frac{9}{a})^h) \quad (*) \\
&= a^h\frac{1-(\frac{9}{a})^{h+1}}{1-(\frac{9}{a})} \quad \text{assume } a \neq 9 \\
&= \frac{a^{h+1} - 9^{h+1}}{a-9} \\
&= \frac{aa^{\log_3 n} - 9n^2}{a-9} \\
&= \frac{an^{\log_3 a} - 9n^2}{a-9}
\end{aligned}$$

When $a = 9$, $T(n) = a^h(h+1) = a^{\log_3 n}(\log_3 n + 1) = n^{\log_3 a}(\log_3 n + 1) = n^2(\log_3 n + 1)$.

2. When $a < 9$,

$$T(n) = \frac{9n^2 - an^{\log_3 a}}{9 - a} = O(n^2).$$

When $a = 9$,

$$T(n) = n^2(\log_3 n + 1) = O(n^2 \log n).$$

When $9 < a \leq 27$,

$$T(n) = \frac{an^{\log_3 a} - 9n^2}{a - 9} = O(n^{\log_3 a}) = O(n^3)$$

Note that when $T(n) = O(n^2)$ is true, $T(n) = O(n^2 \log n)$ is also true; and when $T(n) = O(n^2 \log n)$ is true, $T(n) = O(n^3)$ when $a < 27$ is also true.

So the final answers are:

- $T(n) = O(n^2)$ when $a < 9$
- $T(n) = O(n^2 \log n)$ when $a \leq 9$
- $T(n) = O(n^3)$ when $a \leq 27$

Grading: 5, 6.
For Part 1, 3 points for reaching (*), 1 point for a formula that involves only n, 1 for separating the case $a = 9$.

For Part 2, 2 points for each sub-question. Note that $a < 9$ is also correct for (ii), but the bound is not tight. So give 1. For (iii), $a \leq 9$ is also correct, but the bound is not tight. Give. 1

Problem 3: [10 points] Let $n \geq 1$ be an integer that is a power of 2. Consider a function $T(n)$ such that $T(1) = 1$ and, for $n > 1$,

$$T(n) \leq 8T(n/2) + 3n^2 + 2n + 7.$$

Prove that $T(n) = O(n^3)$.

Solution: It suffices to show that there exist constants n_0, $c_1 > 0$ and $c_2 > 0$ such that $T(n) \leq c_1 n^3 - c_2 n^2$ for all $n > n_0$.

Set $n_0 = 0$.
Base case ($n = 1$): $T(1) = 1 \leq c_1 1^3 - c_2 1^2 = c_1 - c_2$. This is true when $c_1 \geq c_2 + 1$.

Assume $T(n) \leq c_1 n^3 - c_2 n^2$ is true for $n = 2^i$ where $i = 0, \cdots, j-1$. For $n = 2^j$,

$$\begin{aligned}
T(n) &\leq 8T(n/2) + 3n^2 + 2n + 7 \\
&\leq 8(c_1 n^3/8 - c_2 n^2/4) + 3n^2 + 6n + 7 \\
&= c_1 n^3 - c_2 n^2 - c_2 n^2 + 3n^2 + 6n + 7 \\
&\leq c_1 n^3 - c_2 n^2 - c_2 n^2 + 3n^2 + 6n^2 + 7n^2 \\
&= c_1 n^3 - c_2 n^2 + (16 - c_2) n^2 \\
&\leq c_1 n^3 - c_2 n^2, \qquad \text{when } 16 - c_2 \leq 0.
\end{aligned}$$

By the principle of MI, we have proved that $T(n) \leq c_1 n^3 - c_2 n^2$ when $n_0 = 0, c_2 \geq 16$ and $c_1 \geq c_2 + 1$. Thus, $T(n) = O(n^3)$.

Grading: 3 points if stronger inductive hypothesis is attempted, 2 points for the base case, 2 points if correct logic is employed for specifying the constraints, 3 points if the rest is correct.

Problem 4: [10 points] This question involves two bags of balls and two players. The first bag contains 1 red ball and 9 blue balls and the second bag contains 9 red balls and 1 blue ball. The two players are Tom and Jerry. They each pick one ball from one of the bags. Let E_t be the event that the ball picked by Tom is red and E_j be the event that the ball picked by Jerry is red.

1. Tom picks one ball from the first bag and puts it back. Then Jerry picks one ball from the same bag. What are $P(E_t)$, $P(E_j)$, and $P(E_j|E_t)$?
2. Tom randomly chooses one bag, picks one ball from that bag and puts it back. Then Jerry picks one ball from the same bag. What are $P(E_t)$, $P(E_j)$, and $P(E_j|E_t)$?

Explain how you obtain your answers. Note that if we know the bag from which Tom and Jerry pick the balls, E_t and E_j are independent. The independence is not true if we do not know the bag.

Solution:

1. Because there are 1 red ball and 9 blue balls in the first bag, $P(E_t) = \frac{1}{10}$. Because Tom puts his ball back to the bag, there are also 1 red ball and 9 blue balls in the first bag when Jerry picks his ball. So, $P(E_j) = \frac{1}{10}$. $P(E_j|E_t) = P(E_j \cap E_t)/P(E_t) = P(E_j)P(E_t)/P(E_t) = \frac{1}{10} \times \frac{1}{10}/\frac{1}{10} = \frac{1}{10}$.

2. Let B_1 be the event that Tom picks the first bag, and B_2 be the event that Tom picks the seond bag. We have

$$P(E_t) = P(B_1)P(E_t|B_1) + P(B_2)P(E_t|B_2) = \frac{1}{2} \times \frac{1}{10} + \frac{1}{2} \times \frac{9}{10} = \frac{1}{2}.$$

Because Tom puts his ball back to the bag, Jerry faces exactly the same situation as Tom. So, $P(E_j) = \frac{1}{2}$.

Next consider $P(E_j \cap E_t)$:

$$\begin{aligned} P(E_j \cap E_t) &= P(B_1)P(E_j \cap E_t|B_1) + P(B_2)P(E_j \cap E_t|B_2) \\ &= P(B_1)P(E_j|B_1)P(E_t|B_1) + P(B_2)P(E_j|B_2)P(E_t|B_2) \\ &= \frac{1}{2} \times \frac{1}{10} \times \frac{1}{10} + \frac{1}{2} \times \frac{9}{10} \times \frac{9}{10} \\ &= \frac{82}{200}. \end{aligned}$$

So,

$P(E_j|E_t) = P(E_j \cap E_t)/P(E_t) = \frac{82}{200}/\frac{1}{2} = \frac{82}{100}$.

It is interesting to note that, in this case, $P(E_j|E_t) \neq P(E_j)$.

Grading: 4 (1, 1 2); 6 (1.5, 1.5, 3) .

Problem 5: [12 points] The professor of a class of n students is lazy and lets his students to grade their own homework. After the answer sheets are collected, he hands them back to the students randomly for grading. Each student gets one homework to grade.

Tom, Jerry and Spike are three students in the class and they are friends. Let E_{tt} be the event that Tom grades his own homework, and E_{t3} be the event that Tom grades the homework of one of the three friends (i.e., Tom, Jerry and Spike). The events E_{jj}, E_{ss}, E_{j3} and E_{s3} are defined similarly.

1. What is $P(E_{tt})$?
2. What is $P(E_{tt} \cup E_{jj} \cup E_{ss})$?
3. What is $P(E_{t3})$?
4. What is $P(E_{t3} \cup E_{j3} \cup E_{s3})$?

Explain how you obtain your answers.

Solution:

1. Similar to the backpack problem, a distribution of answer sheets to the students (outcome) is a permutation of $[1, 2, ..., n]$. Among all the $n!$ equally-likely outcomes, Tom gets his own sheet in $(n-1)!$ of them. Therefore $P(E_{tt}) = \frac{(n-1)!}{n!} = \frac{1}{n}$

2. Using the same reasoning in part (1):

   $P(E_{tt}) = P(E_{jj}) = P(E_{ss}) = \frac{1}{n}$
   $P(E_{tt} \cap E_{jj}) = P(E_{tt} \cap E_{ss}) = P(E_{jj} \cap E_{ss}) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$
   $P(E_{tt} \cap E_{jj} \cap E_{ss}) = \frac{(n-3)!}{n!} = \frac{1}{n(n-1)(n-2)}$

   Then, by the principle of inclusion and exclusion,

$$
\begin{aligned}
P(E_{tt} \cup E_{jj} \cup E_{ss}) &= P(E_{tt}) + P(E_{jj}) + P(E_{ss}) \\
&\quad -P(E_{tt} \cap E_{jj}) - P(E_{tt} \cap E_{ss}) - P(E_{jj} \cap E_{ss}) \\
&\quad +P(E_{tt} \cap E_{jj} \cap E_{ss}) \\
&= \frac{3}{n} - \frac{3}{n(n-1)} + \frac{1}{n(n-1)(n-2)} \\
&= \frac{3(n-1)(n-2) - 3(n-2) + 1}{n(n-1)(n-2)} \\
&= \frac{3n^2 - 12n + 13}{n(n-1)(n-2)}
\end{aligned}
$$

3. There are $(n-1)!$ outcomes where Tom gets his own homework. (Part (1))
Similarly, there are $(n-1)!$ outcomes where Tom gets Jerry's homework.
There are also $(n-1)!$ outcomes where Tom gets Spike's homework.
In total, there are $3(n-1)!$ outcomes where Tom gets the homework of one of the three friends.
Hence, $P(E_{t3}) = \frac{3(n-1)!}{n!} = \frac{3}{n}$

4. $P(E_{t3}) = P(E_{j3}) = P(E_{s3}) = \frac{3}{n}$

   Consider $E_{t3} \cap E_{j3}$. This is the event where Tom and Jerry both get any two of the three friends' homework copies. There are $3 \cdot 2 = 6$ different possibilities. For each possibility, there are $(n-2)!$ corresponding outcomes. Therefore, the total number of outcomes for this event is $6(n-2)!$.
   $P(E_{t3} \cap E_{j3}) = P(E_{t3} \cap E_{s3}) = P(E_{j3} \cap E_{s3}) = \frac{6(n-2)!}{n!} = \frac{6}{n(n-1)}$

   Consider $E_{t3} \cap E_{j3} \cap E_{s3}$. This is the event where Tom, Jerry, and Spike all get one of the three friends' homework copies. There are $3! = 6$ different possibilities. For each possibility, there are $(n-3)!$ corresponding outcomes. Therefore, the total number of outcomes for this event is $6(n-3)!$.
   $P(E_{t3} \cap E_{j3} \cap E_{s3}) = \frac{6(n-3)!}{n!} = \frac{6}{n(n-1)(n-2)}$

   Then, by the principle of inclusion and exclusion,

$$\begin{aligned}
P(E_{t3} \cup E_{j3} \cup E_{s3}) &= P(E_{t3}) + P(E_{j3}) + P(E_{s3}) \\
&\quad -P(E_{t3} \cap E_{j3}) - P(E_{t3} \cap E_{s3}) - P(E_{j3} \cap E_{s3}) \\
&\quad +P(E_{t3} \cap E_{j3} \cap E_{s3}) \\
&= \frac{9}{n} - \frac{18}{n(n-1)} + \frac{6}{n(n-1)(n-2)} \\
&= \frac{9(n-1)(n-2) - 18(n-2) + 6}{n(n-1)(n-2)} \\
&= \frac{3(3n^2 - 15n + 20)}{n(n-1)(n-2)}
\end{aligned}$$

Grading: 2, 3, 2, 5 .

Student ID: ______________

Problem 6: [10 points] Consider throwing m balls into n boxes. The probability of each ball ending up in any given box is $\frac{1}{n}$. Let X be the number of balls in the first box and Y be the number of boxes that are empty.

1. What is $E(X)$?
2. What is $V(X)$?
3. What is $E(Y)$?

Show the steps of your calculations.

Solution:

1. $E(X)$: This is a Bernoulli trials process with m trials and probability $\frac{1}{n}$ of success. Thus, by Theorem 5.12, $E(X) = m\left(\frac{1}{n}\right) = \frac{m}{n}$

   Alternative solution:
   Let X_i be the indicator random variable such that $X_i = 1$ if the i-th ball is in the first box, $X_i = 0$ otherwise. Obviously, $X = \sum_{i=1}^{m} X_i$. For $1 \leq i \leq m$, $E(X_i) = P(\text{ball } i \text{ in the first box}) = \frac{1}{n}$.
   Therefore,

$$\begin{aligned} E(X) &= E(\sum_{i=1}^{m} X_i) \\ &= \sum_{i=1}^{m} E(X_i) \\ &= \sum_{i=1}^{m} \frac{1}{n} \\ &= \frac{m}{n} \end{aligned}$$

2. $V(X)$: This is a Bernoulli trials process with m trials and probability $\frac{1}{n}$ of success. Thus, by Theorem 5.X (or L18:Theorem #2 for L2), $V(X) = m\left(\frac{1}{n}\right)\left(1 - \frac{1}{n}\right) = \frac{m(n-1)}{n^2}$.

   Alternative solution:
   $V(X) = V(\sum_{i=1}^{m} X_i) = \sum_{i=1}^{m} V(X_i)$ since X_i and X_j are independent for $i \neq j$.
   For $1 \leq i \leq m$, $V(X_i) = (0 - E(X_i))^2\left(\frac{n-1}{n}\right) + (1 - E(X_i))^2\left(\frac{1}{n}\right) = \frac{n-1}{n^3} + \frac{(n-1)^2}{n^3} = \frac{n(n-1)}{n^3} = \frac{n-1}{n^2}$
   Therefore, $V(X) = \sum_{i=1}^{m} \frac{n-1}{n^2} = \frac{m(n-1)}{n^2}$

3. $E(Y)$: Let Y_i be the indicator random variable such that $Y_i = 1$ if box i is empty, $Y_i = 0$ otherwise. Obviously, $Y = \sum_{i=1}^{n} Y_i$.
For $1 \leq i \leq n$, $E(Y_i) = P(\text{box } i \text{ is empty}) = \left(\frac{n-1}{n}\right)^m$.
Therefore,

$$\begin{aligned} E(Y) &= E(\sum_{i=1}^{n} Y_i) \\ &= \sum_{i=1}^{n} E(Y_i) \\ &= \sum_{i=1}^{n} \left(\frac{n-1}{n}\right)^m \\ &= n \left(\frac{n-1}{n}\right)^m \end{aligned}$$

Grading: 3, 3, 4.

Student ID: ______________

Problem 7: [15 points] Consider a student who does an online practice test. He starts by getting a set of 10 problems from the computer. If he can solve at least 9 of the 10 problems, he stops. Otherwise, he continues by getting another set of 10 problems, and so on. Let X be the total number of **problem sets** that he works on.

There are two scenarios regarding the probability p that he can solve each problem: (1) $p = 0.8$ and does not change over time, and (2) $p = 0.8$ for the first problem set, $p = 0.9$ for the second problem set, and $p = 1.0$ thereafter.

1. What is $P(X = 1)$ in Scenario 1? Show the steps of calculation.
2. What is $E(X)$ in Scenario 1? Explain your answer.
3. What is $E(X)$ in Scenario 2? Show the steps of calculation.
4. In which scenario is $V(X)$ larger? Answer this question based on your intuition. There is **no** need to calculate $V(X)$.

Solution:

1. Test binomial distribution
   $X = 1$ means the student worked only 1 problem set, and then stopped. He needs to solve at least 9 of 10 problems in this first problem set. That is to say, he needs to solve 9 of 10 problems or solve all the 10 problems. Therefore,
   $P(X = 1) = \binom{10}{9} p^9(1 - p) + p^{10} = 10 \times 0.8^9 \times 0.2 + 0.8^{10} = 0.3758$

2. Test expected time until first success
   X models the number of sets of problems (trials) the student tries (performs) until the first success, so $E(X) = \frac{1}{\text{probability of success}}$.
   In Scenario 1, the *probability of success* (i.e., solve at least 9 out of 10 problems) does not change. Therefore, it is equal to
   $P(X = 1) = 0.3758$.
   Consequently, $E(X) = \frac{1}{\text{probability of success}} = \frac{1}{0.3758} = 2.6420$

3. Test brute-force calculation of expectation
   In Scenario 2, the possible values for X are $\{1, 2, 3\}$.
   Let Y_i be the event that the student can solve at least 9 of the 10 problems in i^{th} problem set.
   $P(Y_1) = P(X = 1) = 0.3758$
   $P(Y_2) = \binom{10}{9} (0.9)^9(1 - 0.9) + (0.9)^{10} = 10 \times 0.9^9 \times 0.1 + 0.9^{10} = 0.7362$
   $P(Y_3) = 1$
   And then, we know $P(X = 1) = 0.3758$

$P(X=2) = P(Y_2)P(\bar{Y}_1) = 0.7362 \times (1 - 0.3758) = 0.4595$
$P(X=3) = P(Y_3)P(\bar{Y}_2)P(\bar{Y}_1) = 1 \times (1 - 0.7362) \times (1 - 0.3758) = 0.1647$
Therefore, $E(X) = 1 \times 0.3758 + 2 \times 0.4595 + 3 \times 0.1647 = 1.7889$

4. Answer: Scenario 1. Test intuitive understanding of variance.

Grading: 3, 3, 6, 3.

Student ID: ______________

Problem 8: [10 points] Let n be a positive integer. Give a combinatorial proof of the following identity,

$$\binom{2n}{2} = 2\binom{n}{2} + n^2$$

An algebraic proof of this identity will **not** be accepted as a solution.

Solution:

- Left side: select two items from $2n$ items.
- Right side: split $2n$ items into two parts, each containing n items. We have $\binom{n}{2}$ ways to select two items from the first part, and $\binom{n}{2}$ ways to select two items from the second part. Moreover, there are n^2 ways to select two items such that one is in the first part and the other in the second part. Observe that $2\binom{n}{2} + n^2$ is total the number of ways to select two items from $2n$ items (all possible cases: (i) the two items are in the first part, (ii) the two items are in the second part, (iii) one item is in the first part and the other in the second).

Both left and right hand side are counting the number of ways to select two items from $2n$ items, so they are equal.

Grading: 2 points for the left side, 3 points for attempting to split the $2n$ items, and 5 points for a correct explanation for the right side.

Student ID: ______________

Problem 9: [12 points] Let a, e and n be three positive integers.

1. Describe the repeated squaring method for computing $a^e \bmod n$.
2. Let $T(e)$ be the number of multiplications carried out in repeated squaring. Prove that $T(e) = O(\log e)$.

Solution:

1. Algorithm

   (a) Find $0 \le k_1 < k_2 \le \ldots \le k_t$, such that $e = 2^{k_1} + 2^{k_2} + \ldots + 2^{k_t}$

   - To do this, first obtain the binary representation of e.
   - Then scan the binary representation from right to left.
   - The k_i are just the locations of 1's.
   - Example: $e = 50$. Binary representation: 110010. The k_i's: $k_1 = 1, k_2 = 4, k_3 = 5$.

   (b) Calculate: $I_0 = a$ and, for all $i = 1, 2, \ldots, k_t$, $I_i = (I_{i-1})^2 \bmod n$

   (c) Calculate: $a^e \bmod n = (((I_{k_1} \cdot I_{k_2}) \bmod n) \cdots I_{k_t}) \bmod n$

2. Complexity

   - At Step 2, k_t multipliations are performed.
   - At Step 3, no more the k_t multiplication are performed.
   - So, $T(e) \le 2k_t$. (*)
   - On the other hand,

   $$e = 2^{k_1} + 2^{k_2} + \ldots + 2^{k_t} \ge 2^{k_t}$$

   - so, $k_t \le \log_2 e$. (**)
   - Hence $T(e) \le 2k_t \le 2\log_2 e$.
   - Therefore, $T(e) = O(\log e)$. (***)

Grading: Algorithm: 6 points, 2 for each step, ok if the method for finding k_i's is not described. Complexity analysis: 6 points, 2 for (*), 2 for (**), 2 for (***)